

Appl. No. 10/019,882  
Amdt. Dated July 2, 2007  
Reply to Office Action of April 20, 2007

**Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (previously presented) A method comprising:
  - (a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;
  - (b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and
  - (c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model.
2. (original) The method of claim 1, wherein parts (a) – (c) are repeated at least once.
3. (original) The method of claim 1, wherein the utterances are converted into a recognized phone string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.
4. (previously presented) The method of claim 1, wherein calculating the estimated weights comprises computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.
5. (previously presented) The method of claim 1, wherein calculating the estimated weights comprises computing an average likelihood difference per frame according to equation (1) as follows:

$$L_n = \frac{H_e^n - H_b^n}{H_e^n - H_b^n} - \frac{R_e^n - R_b^n}{R_e^n - R_b^n} \quad (1),$$

Appl. No. 10/019,882  
 Amdt. Dated July 2, 2007  
 Reply to Office Action of April 20, 2007

where  $H_L^n$  is a log likelihood of hypothesis word  $n$ ,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

6. (previously presented) The method of claim 5, wherein calculating the estimated weights further comprising a weight for misrecognized words of a particular speaker "i" according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

7. (previously presented) The method of claim 1, wherein for a particular speaker, different misrecognized words have different weights.

8. (previously presented) A method comprising:

- (a) recognizing utterances through converting the utterances into a recognized string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model.

9. (previously presented) The method of claim 8, wherein the utterances are converted into the recognized string through applying the speaker independent model.

10. (original) The method of claim 8, wherein parts (b) – (e) are repeated until differences between the reference and recognized strings are less than a threshold.

Appl. No. 10/019,882  
 Amdt. Dated July 2, 2007  
 Reply to Office Action of April 20, 2007

11. (previously presented) The method of claim 8, wherein the utterances are converted into a recognized string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.

12. (previously presented) The method of claim 8, wherein calculating the estimated weights comprises computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over all error words.

13. (previously presented) The method of claim 8, wherein calculating the estimated weights comprises calculating an average likelihood difference per frame according to equation (1) as follows:

$$L_n = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

14. (previously presented) The method of claim 13, wherein calculating the estimated weights comprises calculating a weight for misrecognized words of a particular speaker "i" is calculated according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

15. (previously presented) The method of claim 8, wherein for a particular speaker, different misrecognized words have different weights.

16. (previously presented) An article of manufacture comprising:  
 a storage medium having instructions thereon which when executed cause a processor to perform operations comprising:

Appl. No. 10/019,882  
 Amdt. Dated July 2, 2007  
 Reply to Office Action of April 20, 2007

(a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;

(b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and

(c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model.

17. (previously presented) The article of manufacture of claim 16, wherein parts (a) – (c) are repeated at least once.

18. (previously presented) The article of manufacture of claim 16, wherein the utterances are converted into a recognized phone string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.

19. (previously presented) The article of manufacture of claim 16, wherein the estimated weights are computed through computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.

20. (previously presented) The article of manufacture of claim 16, wherein an average likelihood difference per frame is used to calculate the estimated weights and is computed according to equation (1) as follows:

$$L_n = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

Appl. No. 10/019,882  
Amdt. Dated July 2, 2007  
Reply to Office Action of April 20, 2007

21. (previously presented) The article of manufacture of claim 20, wherein a weight for misrecognized words of a particular speaker "i" is calculated according to equation (2) as follows:

$$w_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

a number of misrecognized words.

22. (previously presented) The article of manufacture of claim 16, wherein for a particular speaker, different misrecognized words have different weights.

23. (previously presented) An article of manufacture comprising:  
a storage medium having instructions thereon which when executed cause a processor to perform operations comprising:

- (a) recognizing utterances through converting the utterances into a recognized phone string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model.

24. (previously presented) The article of manufacture of claim 23, wherein the utterances are converted into the recognized string through applying the speaker independent model.

25. (original) The article of manufacture of claim 23, wherein parts (b) – (e) are repeated until differences between the reference and recognized strings are less than a threshold.

26. (previously presented) The article of manufacture of claim 23, wherein the utterances are converted into a recognized string a first time through applying the speaker

Appl. No. 10/019,882  
 Amdt. Dated July 2, 2007  
 Reply to Office Action of April 20, 2007

independent model and thereafter through applying the most recently obtained speaker dependent model.

27. (previously presented) The article of manufacture of claim 23, wherein the estimated weights are computed through computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.

28. (previously presented) The article of manufacture of claim 23, wherein an average likelihood difference per frame is used to calculate the estimated weights and is calculated according to the equation (1) as follows:

$$L_n = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

29. (previously presented) The article of manufacture of claim 28, wherein a weight for misrecognized words of a particular speaker "i" is calculated according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

30. (previously presented) The article of manufacture of claim 23, wherein for a particular speaker, different misrecognized words have different weights.